EmeraldApp: An AI-Driven Tool for Detecting Greenwashing in Sustainability Claims

Georgios Kaoukis*
g.kaoukis@athenarc.gr
Archimedes, Athena Research Center
Athens, Greece

Ioannis-Aris Koufopoulos*
j.koufopoulos@athenarc.gr
Archimedes, Athena Research Center
Athens, Greece

Eleni Psaroudaki
h.psaroudaki@athenarc.gr
Archimedes, Athena Research Center
Athens, Greece

Danae Pla Karidi†
danae@athenarc.gr
Archimedes, Athena Research Center
Athens, Greece

Evaggelia Pitoura
pitoura@uoi.gr
University of Ioannina
Ioannina, Greece
Archimedes, Athena Research Center
Athens, Greece

George Papastefanatos
gpapas@athenarc.gr
IMSI, Athena Research Center
Athens, Greece

Panayiotis Tsaparas
tsap@uoi.gr
University of Ioannina
Ioannina, Greece
Archimedes, Athena Research Center
Athens, Greece

Abstract

Greenwashing, i.e., presenting misleading environmental claims, undermines the trustworthiness of sustainability reporting. We present EmeraldApp, a tool designed to verify corporate sustainability claims. EmeraldApp classifies claims as greenwashing, not greenwashing, or abstains, and generates fact-based justifications. The tool supports different LLM pipelines: `EM-RAG`, which retrieves evidence from *EmeraldDB*, a vector database of ESG (Environmental, Social, Governance) report chunks; `EM-KGRAG`, which retrieves evidence from *EmeraldGraph*, a domain-specific knowledge graph; `EM-HYBRID`, which combines the other two retrieval pipelines; and `EM-NR`, without retrieval capabilities. EmeraldApp supports quick claim checks for end users, while advanced configuration enables pipeline comparison and deeper inspection of factual evidence and justification. Previously checked claims are stored in *ClaimsDB* and are accessible through the EmeraldApp to improve auditability and transparency.

CCS Concepts

• **Computing methodologies** → **Knowledge representation and reasoning**; • **Information systems** → *Information retrieval*.

Keywords

Greenwashing Detection, Retrieval Augmented Generation (RAG), Knowledge Graph, Responsible AI, Evidence-based Explanations

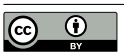
ACM Reference Format:
Georgios Kaoukis, Ioannis-Aris Koufopoulos, Eleni Psaroudaki, Danae Pla Karidi, Evaggelia Pitoura, George Papastefanatos, and Panayiotis Tsaparas. 2026. EmeraldApp: An AI-Driven Tool for Detecting Greenwashing in Sustainability Claims. In *Companion Proceedings of the ACM Web Conference 2026 (WWW Companion '26), April 13–17, 2026, Dubai, United Arab Emirates.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3774905.3794644

1 Introduction

Greenwashing, the deceptive practice of presenting misleading claims about an organization's environmental performance, has become a pervasive obstacle to genuine sustainability progress [2], creating a verification gap: given a corporate sustainability claim (e.g., "We have achieved 100% renewable energy sourcing"), stakeholders often lack the tools to efficiently cross-check such assertions against the vast, dispersed volume of corporate ESG (Environmental, Social, and Governance) disclosures. Manual inspection and verification remain prohibitively labor-intensive and time-consuming [1, 9], while the ambiguity of corporate language makes it difficult to distinguish between verifiable facts and marketing noise.

While large language models (LLMs) excel in general text tasks, they falter in specialized verification due to: (1) hallucination of plausible but false details [3, 4]; (2) absence of domain-specific grounding [10]; and (3) opaque outputs lacking traceable evidence trails. Current solutions like ecoOops [5] and Greenwashing Identifier [7] primarily function as compliance auditors or risk assessment engines, scanning entire corporate reports to assign high-level compliance scores or flag general regulatory discrepancies, with limited

*Both authors contributed equally to this research.
†Corresponding author.



WWW Companion '26, Dubai, United Arab Emirates

ACM ISBN 979-8-4007-2308-7/2026/04
https://doi.org/10.1145/3774905.3794644

support for interactive, evidence-traceable assessment of individual claims.

In contrast, EmeraldApp is an AI-driven tool for automatic greenwashing verification and justification at the level of individual sustainability claims. EmeraldApp is powered by EmeraldMind [8], which introduces multiple RAG-LLM verification pipelines for greenwashing detection. Unlike prior efforts, which mainly provide document-level compliance scores or coarse risk flags, EmeraldApp provides an interactive, evidence-traceable workflow: users submit a claim, run it under multiple verification modes (retrieval-based pipelines `EM-RAG`, `EM-KGRAG`, `EM-HYBRID`, or a no-retrieval mode `EM-NR`), and receive a claim-specific verdict and a short justification. For the RAG pipelines, the justification is grounded in domain-specific evidence stores. These can be in a textual or subgraph format, depending on the retrieval mode. The system also logs verified claims and their justifications, enabling users to search, reopen, compare, and reuse past analyses for auditing and ongoing investigations.

2 EmeraldApp Overview

The EmeraldApp architecture (Figure 1) separates a user-centric frontend from a modular backend that orchestrates claim verification over shared evidence stores. The frontend provides components for claim input and verification configuration, search and inspection of past claims, and a results and visualization area. The backend provides access to several verification pipelines. *EmeraldDB* and *EmeraldGraph* hold ESG evidence in text and graph form, respectively. *ClaimsDB* indexes previously verified claims for search and reuse through the Claim Manager. The source code for EmeraldApp is publicly available at: https://github.com/ai4greenwashing/EmeraldApp.

2.1 Backend

This section describes the core backend components that enable evidence-grounded claim verification in EmeraldApp: the evidence stores that hold ESG knowledge, the pipelines that retrieve and reason over it, and the mechanisms for revisiting past analyses.

2.1.1 Evidence Stores. The backbone of EmeraldApp is its evidence stores for claim verification. *EmeraldDB* is a vectorized document store. It contains ESG report passages as embeddings, along with contextual metadata (e.g., report, company, year, page) and provenance needed for traceable retrieval. *EmeraldGraph* is a domain-specific property graph. It contains ESG entities and typed relations that link organizations to KPI observations, goals, claims, and supporting documents. These stores surface ESG-specific evidence that is often missing from generic knowledge bases. They also enable auditable reasoning by combining retrievable text context with explicit graph paths.

The stores are populated automatically from corporate ESG reports using a parsing and normalization pipeline. KPI definitions [6] from standards such as GRI provide consistent grounding for indicator names, units, and numerical interpretation. A predefined schema (see details in [8]) constrains entity types, relations, and attributes and separates conceptually distinct disclosures (e.g., targets vs. realized performance). This structure supports fine-grained retrieval of evidence (or counter-evidence) for a claim and enables transparent, evidence-backed justifications.

Finally, the annotated sustainability claims form a dedicated *ClaimsDB*, which stores natural language assertions together with their labels, provenance, and, when available, human- or model-generated justifications. This repository serves as a source of linguistic templates for typical corporate statements and as a benchmark for evaluating new predictions, enabling EmeraldApp to support activities such as searching past decisions, reusing verified evidence, and comparing system outputs across pipelines.

2.1.2 Claim Verification Pipelines. EmeraldApp supports four modes for verifying sustainability claims: `EM-RAG`, `EM-KGRAG`, `EM-HYBRID`, and `EM-NR`, implemented according to the `EmeraldMind` framework [8]. After typing a claim, a user can select one of these pipelines, which determines how evidence is retrieved and how the final verdict is produced. `EM-RAG` is the default configuration.

The `EM-RAG` pipeline focuses on retrieving textual evidence from ESG reports stored in *EmeraldDB*. Given a user-claim and the company name, the *Document Retriever* queries the *EmeraldDB* for semantically relevant text chunks from the corresponding ESG reports. The claim and the retrieved textual evidence are then passed to an *LLM Verifier*, which reasons over this context to generate a verdict and justification returned to the user.

The `EM-KGRAG` pipeline leverages the domain-specific *EmeraldGraph* to provide structured factual context. For a given claim, the system locates the relevant entities in *EmeraldGraph*, and *Graph Retriever* uses a graph-based context retrieval algorithm to extract an evidence subgraph. This subgraph, together with the claim, is passed as structured context to the *LLM Verifier*, which produces a verdict and justification analogous to the `EM-RAG` case.

To maximize classification accuracy and coverage, the `EM-HYBRID` pipeline combines both text and graph retrieval, following the hybrid design introduced in EmeraldMind [8]. When a user selects `EM-HYBRID`, the system runs the `EM-RAG` and `EM-KGRAG` pipelines in parallel and forwards their verdicts and justifications to an *LLM Judge*. The *LLM Judge* compares the alternatives and returns a final, consolidated verdict and justification to the user, together with the corresponding evidence.

The `EM-NR` pipeline serves as a no-retrieval baseline. The user claim is passed directly to the *LLM Verifier*, which produces a verdict and justification based only on the prompt instructions and the model's internal knowledge. This mode is used to contrast retrieval-grounded verification with unguided LLM reasoning.

2.1.3 Past Claims Management. The *Claims Manager* provides a persistent layer for logging, indexing, and reusing previously verified sustainability claims within EmeraldApp. It stores each claim together with its final verdict, justification, associated evidence (text passages or subgraphs), and metadata such as company, year, and pipeline used. This enables users to search and filter historical decisions, quickly retrieve past analyses for similar or recurring statements, and compare how different pipelines behaved on the same or related claims. By turning past verifications into a queryable knowledge base, EmeraldApp supports auditability, monitoring of model behavior over time, and interactive exploration of emerging greenwashing patterns across companies and reporting periods.

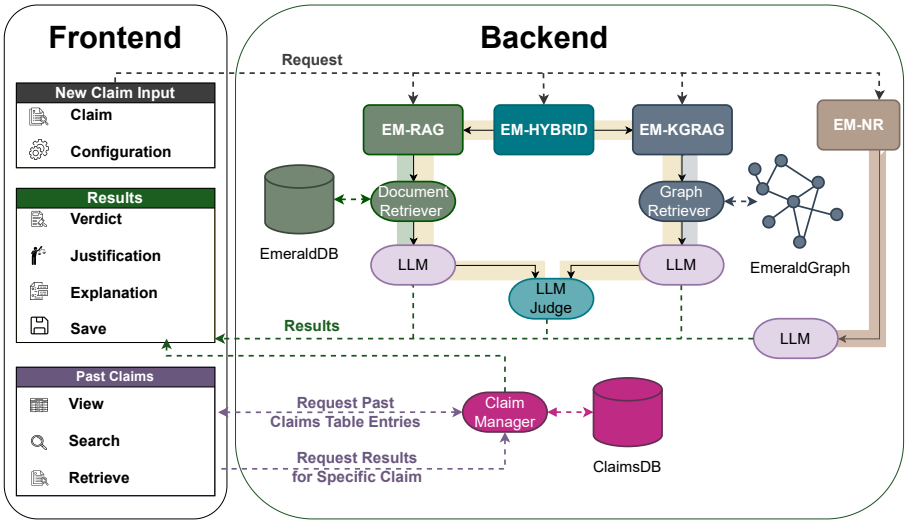


Figure 1: EmeraldApp Overview

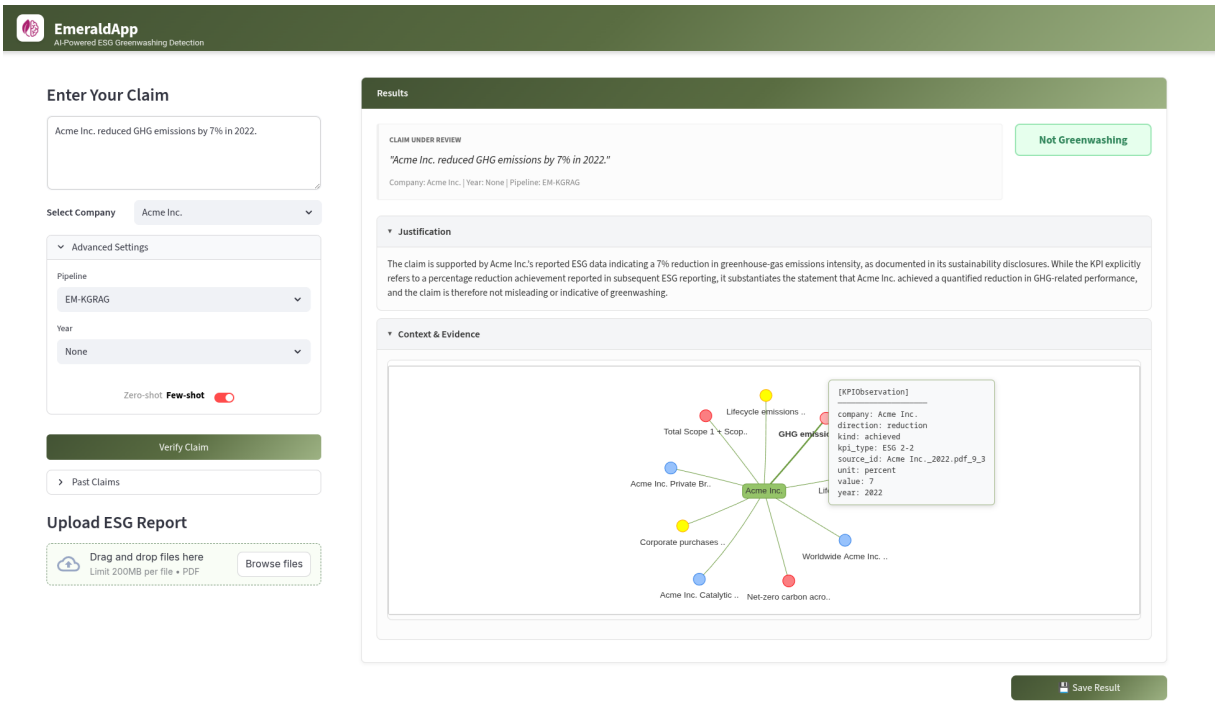


Figure 2: EmeraldApp interface with graph evidence for a non-greenwashing claim

2.2 Graphical User Interface

EmeraldApp's interface follows a dual-panel layout, with the elements on the left dedicated to user input and past claims, and the right reserved for results. The *Enter Your Claim* area on the left lets users type a natural-language sustainability statement and select the company of interest. Optional advanced settings allow users to choose the reporting year, switch pipelines (`EM-RAG`, `EM-KGRAG`, `EM-HYBRID`, or `EM-NR`), and adjust prompting (zero-shot vs. few-shot), as illustrated in Figure 2. A quick query can be performed using the default configuration strategy (`EM-RAG` with a zero-shot prompt) by pressing the *Verify Claim* button without any additional configuration (Figure 3).

The right-hand *Results* panel contains all output for the active claim. A header summarizes the claim under review, the selected company, and the active configuration (year, pipeline, and prompt),

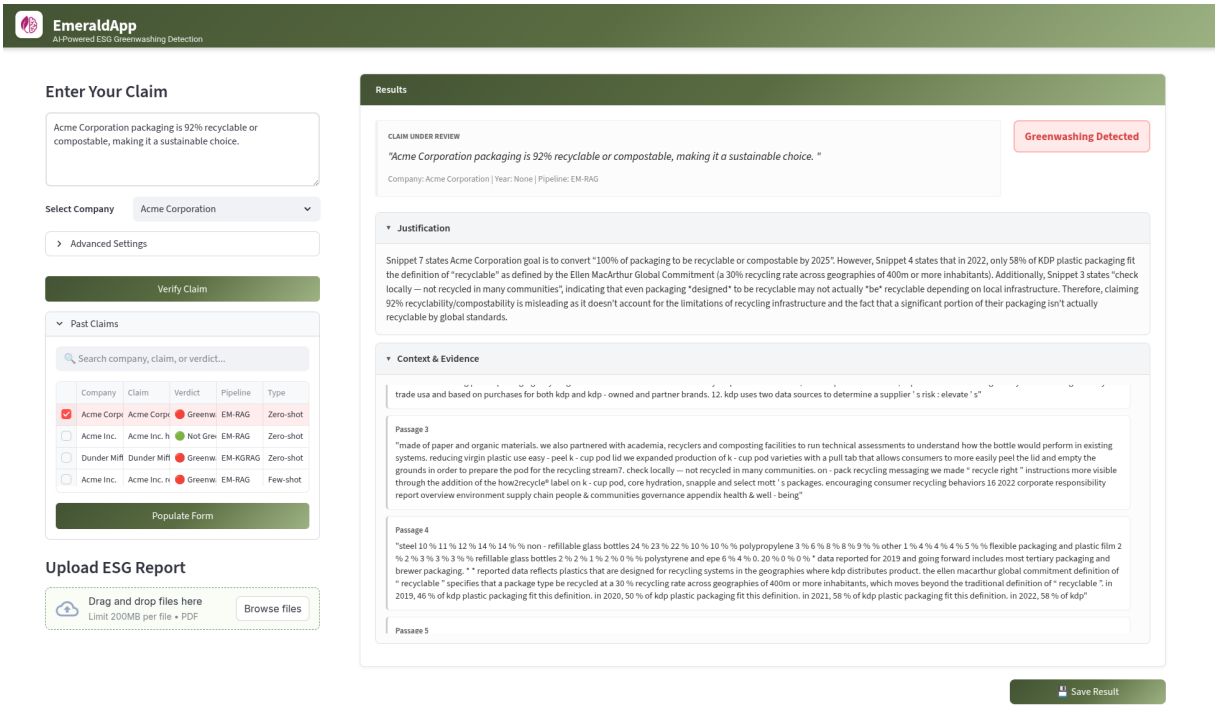


Figure 3: EmeraldApp interface with textual evidence for a detected greenwashing claim

and shows a color-coded verdict badge: red for "Greenwashing Detected" (Figure 3), green for "Not Greenwashing" (Figure 2), and a neutral color for "Abstain". Below, the *Justification* section presents a concise justification grounded in retrieved ESG evidence; for example, in Figure 3, the justification references specific passages from the context. This is essential, as it helps users understand why a verdict was produced rather than treating it as a black-box label.

An expandable *Context & Evidence* view then reveals the underlying support: for text-based pipelines, it displays numbered passages from *EmeraldDB* (Figure 3); for graph-based pipelines, it shows an interactive subgraph from *EmeraldGraph* where hovering over nodes reveals detailed node metadata (Figure 2). These visualizations make the reasoning process tangible for users, allowing them to quickly assess whether the evidence is appropriate and how different configurations change the system's behavior. Finally, a user can save the results of the query made through the save button (Figure 2).

The *Past Claims* collapsible section (Figure 3) includes a table that lists previously verified claims with their companies and verdicts. Search and column-view controls let users filter entries and reveal configuration details. By selecting a row and pressing *Populate Form*, the corresponding claim and settings are reloaded into the *Enter your claim* area, and the associated results are restored in the *Results* section, supporting fast re-runs and deeper evidence inspection. At the bottom, an *Upload ESG Report* widget (Figure 3) supports drag-and-drop or file browsing for PDF reports, which are ingested so that new disclosures become immediately available for verification.

3 Demo Scenarios

During the demo session, the user enters a claim, selects the organization, and runs verification. The system returns a verdict, a short justification, and the most relevant supporting evidence chunks. The user can then open the advanced settings to rerun the same claim with different configurations (e.g., pipeline and prompt) and view the corresponding results and evidence (text, graph). The user visits the past claims menu, filters the claims based on the company of interest, and reopens selected runs. Finally, the user downloads the results with the Save button.

4 Conclusions and Future Work

EmeraldApp is an AI-driven tool for greenwashing detection. It provides a verdict and a fact-based justification for each submitted claim by combining text- and graph-based retrieval from ESG knowledge. As future work, EmeraldApp can be extended with user authentication and personalized workspaces, allowing users to log in, view, and manage their own history of verified claims, selectively save important analyses, and upload and analyze proprietary data. Finally, a dedicated comparison view for running the same claim through multiple pipelines side by side would further help users understand how different configurations affect verdicts, evidence, and justifications.

Acknowledgment

This work has been partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program.

References

[1] Jean-Flavien Bussotti, Luca Ragazzi, Giacomo Frisoni, Gianluca Moro, and Paolo Papotti. 2024. Unknown claims: Generation of fact-checking training examples from unstructured and structured data. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. 12105–12122.

[2] Tom Calamai, Oana Balalau, Théo Le Guenedal, and Fabian M Suchanek. 2025. Corporate Greenwashing Detection in Text–a Survey. *arXiv preprint arXiv:2502.07541* (2025).

[3] Bohan Chen and Andrea L Bertozzi. 2023. AutoKG: Efficient automated knowledge graph generation for language models. In *2023 IEEE International Conference on Big Data (BigData)*. IEEE, 3117–3126.

[4] Hoang Anh Dang, Vu Tran, and Le-Minh Nguyen. 2025. Survey and analysis of hallucinations in large language models: attribution to prompting strategies or model behavior. *Frontiers in Artificial Intelligence* 8 (2025), 1622292.

[5] ecoOops GmbH. 2023. ecoOops: Greenwashing Detector and ESG Analyzer. https://www.ecooops.app/. Accessed: 17 December 2025.

[6] EFFAS and DVFA. 2009. Key Performance Indicators for Environmental, Social & Governance Issues: Financial Analysis and Corporate Valuation. https://ec.europa.eu/docsroom/documents/1547/attachments/1/translations/en/renditions/native Endorsed by EFFAS; Available via European Commission Docsroom.

[7] ImpactScope & Insig AI. 2025. Greenwashing Identifier (GWI). https://gwi.impactscope.com/. Accessed: 17 December 2025.

[8] Georgios Kaoukis, Ioannis-Aris Koufopoulos, Eleni Psaroudaki, Danae Pla Karidi, Evaggelia Pitoura, George Papastefanatos, and Panayiotis Tsaparas. 2026. EmeraldMind: A Knowledge Graph–Augmented Framework for Greenwashing Detection. In *Proceedings of the ACM Web Conference 2026 (WWW '26)*. ACM, Dubai, United Arab Emirates, 11. doi:10.1145/3774904.3792997

[9] Markus Leippold, Saeid Ashraf Vaghefi, Dominik Stammbach, Veruska Muccione, Julia Bingler, Jingwei Ni, Chiara Colesanti Senni, Tobias Wekhof, Tobias Schimanski, Glen Gostlow, et al. 2025. Automated fact-checking of climate claims with large language models. *npj Climate Action* 4, 1 (2025), 17.

[10] Subhey Sadi Rahman, Md Adnanul Islam, Md Mahbub Alam, Musarrat Zeba, Md Abdur Rahman, Sadia Sultana Chowa, Mohaimenul Azam Khan Raiaan, and Sami Azam. 2025. Hallucination to truth: A review of fact-checking and factuality evaluation in large language models. *arXiv preprint arXiv:2508.03860* (2025).